# Impacts of Air Pollution on Chinese Expressions of Happiness on Social Media

**Zixuan Tang[1,*], Xinyi Zhao[1], Juntong Liu[2], Jingshu Li[3], Peiwen Li[4], and Yichen Zhang[5]**

[1]Foreign Studies College, Northeastern University, Shenyang 110819, China
[2]School of Public Administration, Jilin University, Changchun 130012, China
[3]Tourism College, Huaqiao University, Fuzhou 35001, China
[4]Meishi Film Academy, Chongqing University, Chongqing 400044, China
[5]Accounting Studies College, Shanghai University of Finance and Economics Zhejiang College, Jinhua 321000, China
[*]Correspondence: zxuantang@163.com

## Abstract

Many studies have demonstrated that air quality is an important factor affecting national well-being. Is there a relationship between air pollution and people's happiness? To explore this, more than 40,000 haze-related tweets on the Chinese largest microblog platform Sina Weibo were collected. Using daily data for 6 Chinese cities from January 1 to February 12, 2023, we applied natural language processing (NLP) to analyze the Weibo tweets and construct a daily city-level expressed happiness metric. A fixed effects model was applied to reveal the relationship between air pollution and happiness. We found that a one standard deviation rise in the Air Quality Index corresponded to a 0.042 standard deviation fall in the happiness index on average. People in large and rich cities are more sensitive to air pollution, and people suffer more from air pollution on weekends and holidays than on workdays. This project may provide new insights into air pollution and public engagement, and help governments and related institutions to better understand the public's needs regarding air quality.

Keywords:air quality,pollution,NLP,happiness.

# 1 Introduction

## 1.1 Project Background

With large-scale industrialization and extensive use of fossil fuels, air pollution strongly affects the quality of human life. Recent theoretical developments have revealed that heavy air pollution from hazy weather not only poses a great threat to physical health but also exacerbates mental health problems and even leads to depression (Gladka et al., 2022), suicide (Sun et al., 2020), and violent crime (Kuo & Putra, 2021).

Some studies have performed a series of short-term or long-term examinations of the relationship between air quality and people's emotions. However, some of them have focused on a specific group (Wang et al., 2022) or unilateral problem (Hou et al., 2021). Meanwhile, subjective approaches to measuring happiness used in some studies by asking people to access life satisfaction directly, such as questionnaires, are limited to some extent and could be affected by the current situation, the individual's opinion and feelings at the time, or other excessive compounding factors such as weather conditions. In addition to the above limitations, the scope of the investigation in some studies is often either too extensive or too narrow, focusing on one city (F. Zhang et al., 2020) (small-sample instability) or around the globe (vast cultural differences).

Through our study, we can provide new insights into air pollution and public engagement. At least in theory, we arrived at the conclusion that changes in air quality truly affect people's expressed happiness. Based on theoretical aspects, some suggestions are given for control from the perspective of mental health protection. In addition, when people realize that their happiness is closely related to air quality, we hypothesize that they will be more proactive in protecting the environment.

## 1.2 Division of Roles & Responsibilities

For the purpose of enhancing our group's efficiency, we divided the tasks based on our respective subjects and areas of expertise.

1. **Tang Zixuan.** Group Leader. Tang Zixuan was responsible for overall project planning, timeline, and delegation of tasks. She facilitated collaboration and communication, assigned tasks and roles to each team member, scheduled regular team meetings, and provided updates on progress. Besides this, she was in charge of data collection and data analysis.

2. **Liu Juntong.** Researcher. Liu Juntong was responsible for gathering and analyzing relevant information related to the new applications of AI in air pollution and presenting findings to the team clearly and concisely.

3. **Zhao Xinyi.** Questioner. Zhao Xinyi was responsible for constructing the Gantt chart and keeping project progress updated. Meanwhile, she was responsible for ensuring academic rigor, made additions to the research process as well as summarized major findings and implications for practice.

4. **Li Jingshu.** Notetaker. Li Jingshu was responsible for capturing and recording important information and discussions during meetings and summarizing advice or key points to better guide the direction of the research. Also, she worked on the literature review and made class presentations.

5. **Li Peiwen.** Designer. Li Peiwen was responsible for creating visual and graphic elements for projects, creating team introduction videos, and assembling and designing project presentation slides.

6. **Zhang Yichen** Timekeeper. Zhang Yichen was responsible for setting the agenda, monitoring time, and providing reminders. Additionally, she tracked the time spent on various tasks and activities during meetings or work sessions.

## 2 Literature Review

Seligman (2002) distinguished three routes to happiness： living a pleasant life, a good life, or a meaningful life. “Happiness is widely taken to be ‘the object of human desire. . . as being what gives purpose, meaning, and order to human life’”(Ahmed, 2010). In the Collins Dictionary, “HAPPINESS” is defined as a state of well-being characterized by emotions ranging from contentment to intense joy and emotions experienced when in a state of well-being. The Modern Chinese Dictionary explains happiness as "a situation and life that makes people feel good." In other words, happiness is a subjective spiritual emotion - a subjective experience of being happy.

Numerous studies have demonstrated that the environment has a significant impact on individuals’ emotions and that it also affects a wide range of human behaviors. Air pollution triggers a variety of emotions that can pose a threat to society and the public. Moreover, with the same level of air pollution, its influence on cities with higher income, poor air quality, and a higher proportion of women has become greater. But as time has passed, people have moved on from a state of panic and are devoting more attention to policy and the sources of air pollutants (Shi et al., 2023). In this context, an analysis of the correlation between air pollution and people’s expressed happiness on social media may provide a theoretical basis for the formulation of relevant policies. Indeed, the first study evaluating the relationship between particulate matter concentrations and the subjective well-being of its residents was based on data from Ulaanbaatar (Sanduijav et al., 2021).

In China, scholars have explored the correlation between air pollution and residents’ happiness from different perspectives. For example, from an intercity perspective, relevant studies have shown that air pollution, sulfur dioxide, and suspended particulate matter concentrations have negative effects on the individual happiness index(Dang et al., 2020). From the regional perspective, in questionnaire results from the Bohai Rim region, it was found that environmental pollution has a significant negative impact on residents’ happiness(Liu et al., 2022). From the national perspective, drawing on the subjective happiness data of China, researchers examined the effects of seven air pollutants and concluded that PM2.5 and PM10 have a large effect on residents’ happiness (Zhang et al., 2017). International scholars generally believe that air pollution is not conducive to improving people’s happiness. Studies have shown that life satisfaction in Ulaanbaatar is inversely proportional to the level of air pollution, and people have a great willingness to pay for better air quality (Sanduijav et al., 2021). In the case of metropolitan London, life satisfaction (LS) in a selected sample was found to be significantly negatively correlated with subjectively perceived levels of air pollution and measurements of air pollutants at very high spatial resolution (MacKerron & Mourato, 2009). The above analysis of the impact of air pollution on people’s happiness mainly focuses on a certain area, while in the case of a country, results from the the Subjective Wellbeing model (SWB) and the hedonic model lead to the conclusion that low-quality air significantly reduces subjective happiness (Rehdanz & Maddison, 2008).

Although there are also relevant studies on the results of the national air pollution and happiness index in China, there is a relative lack of content focusing on the expressed happiness of social media users. Our goal was to focus on the air quality in China and extensively collect geotagged tweets on Weibo to analyze the relationship between air pollution and people’s happiness.

# 3 Data and Methods

## 3.1 Data Source

To obtain more stable samples, we used a week-by-city pattern to match numerous data sources, such as air quality data, meteorological data, and Weibo tweets.

- **Air pollution data.** Since May 2014, the China National Environmental Monitoring Centre has publicly released hourly air pollutant concentrations (PM2.5, PM10, O3, SO2, CO, and NO2) from monitoring stations covering hundreds of prefecture cities. In that case, the air pollution data of various cities may be easily obtained. Most air quality services in China preferentially display the AQI, which is determined by the maximum concentration of various pollutants and ranges from 0 to 500, with a lower value indicating better air quality. Furthermore, PM2.5 is currently the most prevalent pollutant in China (He et al., 2017). As a result, we selected AQI and PM2.5 as air quality indicators and computed their weekly averages. We collected real-time data points and scaled them down to the city-day level.

- **Weather conditions.** Given that weather and sunlight can influence people's moods, our study uses rich weather measures as controls, such as wind speed, maximum and minimum temperatures, and weather conditions such as rain and clouds. Weather data was obtained from the National Oceanic and Atmospheric Administration (NOAA).

- **Weibo tweets.** Throughout the course of our research, we used web crawler technology to continuously connect to the streaming application programming interface (API) and gathered Weibo posts (updates) from mainland China. The geotagged Weibo data provided through the API were a subset of all Weibo data whose users agreed to contribute their locational information.

## 3.2 Natural Language Processing and Happiness Index

Nowadays, Artificial Intelligence (AI) has greatly impacted the field of Natural Language Processing (NLP) by enabling the development of more advanced and sophisticated NLP models. These AI-powered NLP models are capable of performing a wide range of NLP tasks such as text classification, named entity recognition, machine translation, and sentiment analysis. Sentiment analysis is the task of determining the sentiment or emotion expressed in a piece of text, such as a review, tweet, or blog post. AI has greatly impacted this task and one popular approach in AI-powered sentiment analysis is the use of deep neural networks, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). These models are trained on large datasets of annotated text and learn to identify patterns in the text that are indicative of sentiment. The models can then be used to classify new text into positive, negative, or neutral sentiment categories.

Therefore, based on this AI-powered technology, the Tencent natural language processing (NLP) platform will implement sentiment analysis algorithms to analyze the data collected from Weibo and calculate the sentiment value. The overall happiness index for a city on a given day is calculated by taking the median sentiment value for that city/day. This index has a scale of 0 to 100, with 0 representing a strongly negative mood and 100 representing a strongly positive mood.

### 3.3 Empirical Model

We investigated the effect of air pollution on our happiness index by conducting a fixed effect regression model. The fixed effect model is a variation of the linear regression model. This model helps control for omitted variable bias by modeling differences across groups as fixed and unique to each group, rather than allowing them to vary randomly. This means that the fixed effect model accounts for time-invariant variables and allows for a better estimation of the relationship between the dependent and independent variables. We included the city or area fixed effect, which regulates unobservables that vary across cities, such as temporary traffic restrictions, holidays, and other elements such as seasonal variations in pollution.

We used the fixed effect regression model to identify multiple sentiments found in over 40,000 haze-related tweets on Sina Weibo and evaluate the effect of air quality on people's happiness in 6 Chinese cities (Beijing, Shanghai, Guangzhou, Liaoyang, Zhoukou, and Yuxi) from January 1 to February 12, 2023.

$$\text{HAPPINESS}_{\text{ij}} = \alpha_0 + \alpha_1 \text{ POLLUTION}_{\text{ij}} + \alpha_2 \text{X}_{\text{ij}} + \text{J}_{\text{i}} + \gamma_{\text{i}} \tag{1}$$

$\text{HAPPINESS}_{\text{ij}}$ and $\text{POLLUTION}_{\text{ij}}$ represent city i's happiness index and pollution index on day j, respectively. In the baseline regressions, we used AQI and PM2.5 concentrations as the pollution variables, respectively. $\text{X}_{\text{ij}}$ represents weather conditions, and $\text{J}_{\text{j}}$ was used to control for date fixed effects, where J refers to a set of day indicator variables. We included the city or area fixed effect $\gamma_{\text{i}}$ to account for time-invariant unobservables that differ between cities. Coefficient $\alpha_1$ reflects people's pollution responsiveness, which was expected to be negative.

## 4 Results

In Figure 1, we merged the happiness dataset with the city-level daily AQI, PM2.5 concentration data, and weather data. There is a negative correlation ($P<0.01$, coefficient=0.042, $R^2$=0.61) between PM2.5 concentration and our happiness index. The happiness index ranges from 0 to 100 (where larger values indicate a more positive mood).

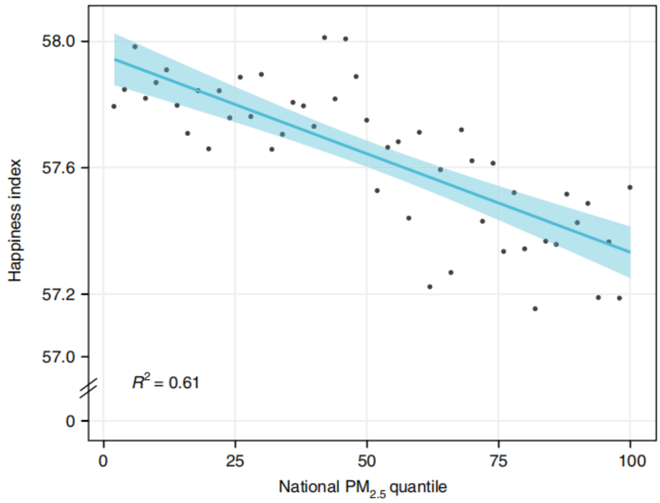


**Figure 1:** The relationship between PM2.5 concentration and the happiness index

The entire city-day sample was divided into 50 groups, each representing a 2 percent range, and sorted by PM2.5 concentration. The graph shows a dot for the median happiness index value for each group, which is fitted by a downward-sloping line. The 95% confidence interval is depicted by the blue shaded region.

Table 1 better illustrates the effect of air pollution on happiness. Column one (All cities) reports our baseline regression results. The coefficients indicate that, on average, a one standard deviation increase in the Air Quality Index is associated with a 0.042 standard deviation decrease in the happiness index.

As shown in column 2, this negative impact grows to twice that in column one (0.089 versus. 0.042), indicating that people in large and rich cities may be more sensitive to air pollution.

**Table 1:** The effect of pollution on expressed happiness

| **Sample** | **All cities** | **Beijing, Shanghai, and Guangzhou** |
|---|---|---|
| AQI | -0.042*** | -0.089*** |
| n | 258 | 129 |
| $R^2$ | 0.579 | 0.718 |

[*] Each column reports the standardized coefficient by employing an OLS estimate equation, controlling for city and date fixed effects, as well as weather variables including temperature, precipitation, wind speed, and the proportion of the day that is cloud-covered. The dependent variable is the median value of the happiness index for all geotagged Weibo posts in a city/day. All 6 cities are included in the sample reported under 'All cities', while the last columns only use data from the three first-tier cities (Beijing, Shanghai, and Guangzhou). Robust standard errors are clustered by city; Pvalues are shown in parentheses, **P<0.05, ***P<0.01.

We also found that when the government declares a "severe pollution" alert, people's happiness significantly decreases according to the happiness index, dropping by 1.21 points. This suggests that people trust the government's announcement and take action accordingly. In response to such alerts, people are more likely to take measures to protect themselves, such as wearing masks or canceling planned activities. This finding has important implications for the government and people's participation in air governance. It indicates that people are likely to participate in efforts to reduce air pollution and improve air quality if they feel that their actions can make a difference. Therefore, the government can encourage people's participation in air governance by providing them with clear information about the sources of air pollution and the actions they can take to reduce it. Additionally, the government can incentivize businesses and industries to reduce their pollution emissions and provide support for the development of cleaner technologies. By working together, the government and the public can take effective steps to improve air quality and protect public health.

# 5 AI and Air pollution

Artificial intelligence (AI) is a new technological science that may be used to develop theories, methods, technologies, and application systems for simulating, extending, and expanding human intelligence. It has been widely used in all levels of environmental governance and will most likely bring about changes to environmental pollution control.

There are three directions in the new applications of AI in air pollution. The first one concerns air pollution forecasting. AI-based techniques are presently considered to be the most transformative technologies for air pollution forecasting due to their specific features such as organic learning, high precision, superior generalization, strong fault tolerance, and better specificity(Guo et al., 2022). Some studies present a comprehensive overview of the most widely used AI-based techniques for air pollution forecasting namely Artificial Neural Networks (ANN), Deep Neural Networks (DNNs), Support vector machines (SVMs) and Fuzzy logic through a systematic literature review (SLR). It is observed that deep neural networks provide better forecasting results(Masood & Ahmad, 2021). The second is urban air pollution monitoring. Artificial intelligence (AI) and machine learning (ML) approaches can be used to build air pollution models to predict pollutant concentrations and assess environmental and health risks. Air pollution data can be uploaded into AI/ML models to estimate different exposure levels within different communities. The most successful global approach is the development of the smart city. However, such an approach can only increase environmental injustice as not all regions have access to AI/ML technologies(Krupnova et al., 2022). The last direction comprises a complete air quality decision support system that includes three subsystems, namely air quality forecast, air quality evaluation, and environmental impact estimation. The system is based on the intersection of air quality forecast and early warning. This system has better prediction performance and provides accurate information and decision support, which shows a good application prospect(Xinyue, 2021).

These articles show that AI-based techniques have triggered a resurgence of interest in air pollution forecasting and offer great potential to fundamentally change the way air pollution is forecasted in the near future(Masood & Ahmad, 2021).

# 6 Conclusion

In the existing literature, a detailed examination of the relationship between environmental pollution and individual happiness with various measurement methods and data sources is to be found. Our research focuses on data sources on people's expressed happiness found in social media. Nowadays, it is more common for people to express their feelings on social media platforms, and the characteristics of people's expressions are also affected by regions and cultures. Therefore, when it comes to studying the relationship between air pollution and people's happiness in a particular area, a more accurate regional division is of greater use. In our research, the main geographical scope was China, so we chose the most popular social media platforms in China as data sources.

Although one's emotions are reflected in expressed happiness on social media, our linguistic measure derived from language processing (NLP) may still contain some noise and is therefore only a partial proxy for the emotions of the average individual in a given city on a given day, and Weibo users' posts serve as the basis for our happiness index. We also recognize that although this group is a sizable representation of Chinese people, it was not chosen at random. Social media users are more apt to be younger and better educated. The elderly, who are less apt to use social media, are actually more susceptible to air pollution. Therefore our results may underestimate the overall negative effect of air pollution on the happiness of a representative sample of the full population.

The future implications of our project are as follows. Firstly, policymakers need to be aware of the influence of public opinion on policies and pay attention to people's release of relevant content on social media when shaping policies and issuing warnings, which is of great significance for their initial implementation and subsequent adjustment. Secondly, according to our study, different economic developments in cities affect people's sensitivity to air pollution, so it is possible to narrow the perceived gaps between urban and rural regions or between the rich and the poor by promoting pollution abatement. While vigorously developing the economy, the government should not neglect the control of air pollution, and should not pursue many economic benefits at the cost of destroying the atmosphere. The harmonious development of people and the environment must be the right way forward. Thirdly, people suffer more air pollution during non-working days, so policymakers should pay attention to the time factor of air pollution control. More precise air pollution control would undoubtedly result in substantial improvements in people's happiness.

**Author Contributions:** Conceptualization: Zixuan Tang, Juntong Liu, Xinyi Zhao, Jingshu Li, Peiwen Li and Yichen Zhang; methodology, validation, formal analysis: Zixuan Tang, Juntong Liu and Xinyi Zhao; investigation, resources, data curation: Jingshu Li, Peiwen Li and Yichen Zhang; writing-original draft preparation, visualization: Zixuan Tang; writing-review and editing, visualization, supervision: Zixuan Tang, Juntong Liu, Xinyi Zhao, Jingshu Li, Peiwen Li and Yichen Zhang; Project administration: Zixuan Tang. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Research Guidelines:** This study has followed the research guidelines of Humanity, Arts and Digital Creativity, Cambridge Academic Programme 2023. Ethical principles were discussed and considered in this study. In our research projects, ethical aspects were considered in the research context, data collection, and data management.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability:** Please contact the corresponding author for all reasonable requests for access to the data.

**Acknowledgments:** We would like to give our thanks to our faculty professors and

teachers at Cambridge University for their guidance.

**Conflicts of Interest:** The authors declare no conflict of interest.



# References

Ahmed, S. (2010). The promise of happiness. In *The promise of happiness.* Duke University Press.

Dang, Y., Chen, L., Zhang, W., Zheng, D., & Zhan, D. (2020). How does growing city size affect residents' happiness in urban china? a case study of the bohai rim area. *Habitat International*, *97*, 102120.

Gladka, A., Zatonski, T., & Rymaszewska, J. (2022). Association between the long-term exposure to air pollution and depression. *ADVANCES IN CLINICAL AND EXPERIMENTAL MEDICINE*.

Guo, Q., Ren, M., Wu, S., Sun, Y., Wang, J., Wang, Q., ... Chen, Y. (2022). Applications of artificial intelligence in the field of air pollution: A bibliometric analysis. *Frontiers in Public Health*, 2972.

He, J., Gong, S., Yu, Y., Yu, L., Wu, L., Mao, H., ... others (2017). Air pollution characteristics and their relation to meteorological conditions during 2014–2015 in major chinese cities. *Environmental pollution*, *223*, 484–496.

Hou, Y., Gao, M., Huang, L., & Wang, Q. (2021, JUN). Air pollution reduces interpersonal trust: The roles of emotion and emotional susceptibility. *INTERNATIONAL JOURNAL OF ENVIRONMENTAL RESEARCH AND PUBLIC HEALTH*, *18*(11). doi: 10.3390/ijerph18115631

Krupnova, T. G., Rakova, O. V., Bondarenko, K. A., & Tretyakova, V. D. (2022). Environmental justice and the use of artificial intelligence in urban air pollution monitoring. *Big Data and Cognitive Computing*, *6*(3), 75.

Kuo, P.-F., & Putra, I. G. B. (2021). Analyzing the relationship between air pollution and various types of crime. *PLOS ONE*, *16*(8).

Liu, Q., Dong, G., Zhang, W., & Li, J. (2022). The influence of air pollution on happiness and willingness to pay for clean air in the bohai rim area of china. *International Journal of Environmental Research and Public Health*, *19*(9).

MacKerron, G., & Mourato, S. (2009). Life satisfaction and air quality in london. *Ecological Economics*, *68*(5), 1441–1453.

Masood, A., & Ahmad, K. (2021). A review on emerging artificial intelligence (ai) techniques for air pollution forecasting: Fundamentals, application and performance. *Journal of Cleaner Production*, *322*, 129072.

Rehdanz, K., & Maddison, D. (2008). Local environmental quality and life-satisfaction in germany. *Ecological economics*, *64*(4), 787–797.

Sanduijav, C., Ferreira, S., Filipski, M., & Hashida, Y. (2021). Air pollution and happiness: Evidence from the coldest capital in the world. *Ecological Economics*, *187*, 107085.

Shi, B., Xu, K., & Zhao, J. (2023). The long-term impacts of air quality on fine-grained online emotional responses to haze pollution in 160 chinese cities. *Science of The Total Environment*, *864*, 161160.

Sun, Y., Chan, T. O., Xie, J., Sun, X., & Huang, Y. (2020, SEP). Examining spatial association of air pollution and suicide rate using spatial regression models. *SUSTAINABILITY*, *12*(18).

Wang, Z. H., Zhao, W. H., Wang, B., Liu, J., Xu, S. L., Zhang, B., ... Guan, D. B. (2022, JUN 17). Environmentally vulnerable or sensitive groups exhibiting varying concerns toward air pollution can drive government response to improve air quality. *ISCIENCE*, *25*(6).

Xinyue, M. (2021). *Research on modeling of air quality decision support system based on artificial intelligence* (Unpublished doctoral dissertation). Lanzhou University.

Zhang, Xin, Chen, Xi, & Xiaobo. (2017). Valuing air quality using happiness data: The case of china. *Ecological economics*.

Zhang, F., Xu, N., Wang, L., & Tan, Q. (2020). The effect of air pollution on the healthy growth of cities: An empirical study of the beijing-tianjin-hebei region. *APPLIED SCIENCES-BASEL*.